

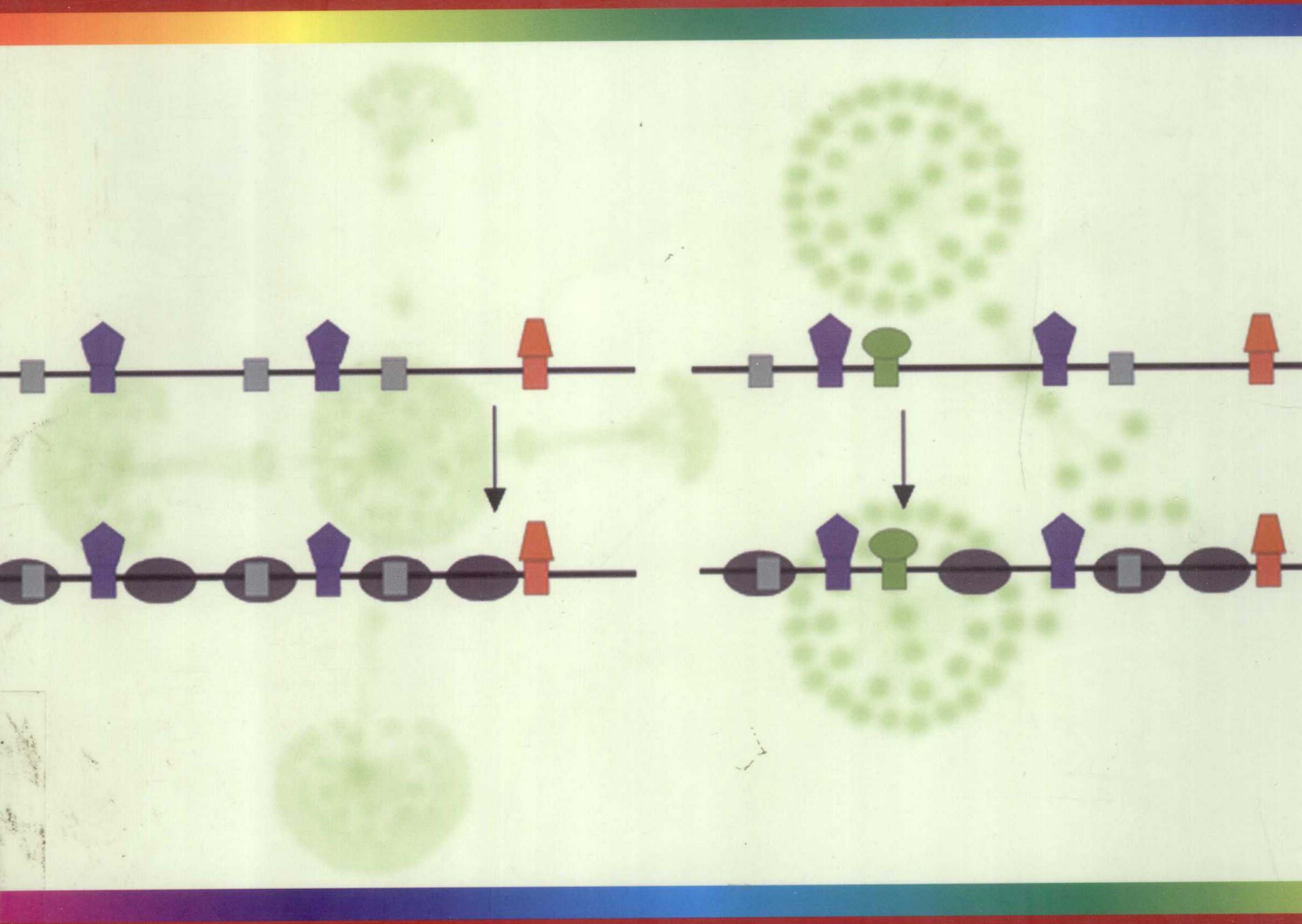
PROCEEDINGS OF THE 7TH ASIA-PACIFIC **BIOINFORMATICS** CONFERENCE

Editors

Michael Q. Zhang

Michael S. Waterman

Xuegong Zhang



Tsinghua University Press

Beijing, China

13-16 January 2009

PROCEEDINGS OF THE 7TH ASIA-PACIFIC
BIOINFORMATICS
CONFERENCE

Editors

Michael Q. Zhang

Cold Spring Harbor Laboratory, USA,
and Tsinghua University, China

Michael S. Waterman

University of Southern California, USA,
and Tsinghua University, China

Xuegong Zhang

Tsinghua University, China

Tsinghua University Press
Beijing

Published by
Tsinghua University Press
Suite A, Xue-Yan Building
Tsinghua University
Beijing 100084
China

PROCEEDINGS OF THE 7TH ASIA-PACIFIC BIOINFORMATICS CONFERENCE

Copyright ©2009 by Tsinghua University Press

The copyright of the papers and abstracts included in these Proceedings are retained by the original authors. This book as a whole may not be reproduced in any form by any means, electronic or mechanical, including photocopying, recording or any information storage and retrieval system now known or to be invented, without written permission from the Publisher. Selected papers from these Proceedings are simultaneously published online as a Supplement Issue of BMC Bioinformatics, for which the BioMed Central Open Access license applies.

版权所有，侵权必究。侵权举报电话：010-62782989 13701121933

图书在版编目（CIP）数据

第7届亚太生物信息学大会论文集=Proceedings of the 7th Asia-Pacific Bioinformatics Conference: 英文/（美）张（Zhang,M.Q.），（美）沃特曼（Waterman,M.S.），张学工主编.
—北京：清华大学出版社，2009.1

ISBN 978-7-302-19048-6

I. 第… II. ①张… ②沃… ③张… III. 生物信息学—国际学术会议—文集—英文
IV. Q811. 4-53

中国版本图书馆 CIP 数据核字（2008）第 190879 号

责任编辑：王一玲

责任印制：李红英

出版发行：清华大学出版社

地 址：北京清华大学学研大厦 A 座

<http://www.tup.com.cn>

邮 编：100084

社 总 机：010-62770175

邮 购：010-62786544

投稿与读者服务：010-62776969, c-service@tup.tsinghua.edu.cn

质 量 反 馈：010-62772015, zhiliang@tup.tsinghua.edu.cn

印 刷 者：清华大学印刷厂

装 订 者：北京市密云县京文制本装订厂

经 销：全国新华书店

开 本：210×297 印 张：60

版 次：2009 年 1 月第 1 版 印 次：2009 年 1 月第 1 次印刷

印 数：1~500

定 价：290.00 元

本书如存在文字不清、漏印、缺页、倒页、脱页等印装质量问题，请与清华大学出版社出版部联系调换。联系电话：(010)62770177 转 3103 产品编号：029820-01

APBC2009

ORGANIZATION

Conference Co-Chairs

- Michael Q. Zhang, Cold Spring Harbor Laboratory and Tsinghua University
- Yanda Li, Tsinghua University
- Xuegong Zhang, Tsinghua University

Organizing Committee

- Xuegong Zhang (Chair), Tsinghua University
- Luhua Lai, Peking University
- Zhirong Sun, Tsinghua University
- Jingchu Luo, Peking University
- Jingdong Jackie Han, Inst of Genetics & Development, CAS
- Shao Li (Poster Co-Chair), Tsinghua University
- Rui Jiang, Tsinghua University
- Xiaoyan Zhu (Poster Co-Chair), Tsinghua University
- Yonglong Zhuang (Publicity Chair), Bioknow Co.

APBC Steering Committee

- Phoebe Chen (Chair), Deakin University
- Sang Yup Lee, KAIST
- Satoru Miyano, University of Tokyo
- Mark Ragan, University of Queensland
- Limsoon Wong, National University of Singapore
- Michael Q. Zhang, Cold Spring Harbor Laboratory and Tsinghua University

Program Committee Co-Chairs

- Michael S. Waterman, University of Southern California and Tsinghua University
- Michael Q. Zhang, Cold Spring Harbor Laboratory and Tsinghua University
- Xuegong Zhang, Tsinghua University

Tutorial Co-Chairs

- Eric P. Xing, Carnegie Mellon University
- Rui Jiang, Tsinghua University

Program Committee Members

- Tatsuya Akutsu, Kyoto University

- Joel Bader, Johns Hopkins University
- Harmen Bussemaker, Columbia University
- Andrea Califano, Columbia University
- Francis Y.L. Chin, Hong Kong University
- Phoebe Chen, Deakin University
- Runsheng Chen, Institute of Biophysics, Chinese Academy of Sciences
- Xue-wen Chen, University of Kansas
- Nadia El-Mabrouk, University of Montreal
- Mikhail Gelfand, Russian Academy of Science
- Roderic Guigo, Centre de Regulacio Genomica
- Jing-Dong Jackie Han, Institute of Genetics & Development, Chinese Academy of Sciences
- Jenn-Kang Hwang, National Chiao Tung University
- Sridhar Hannenhalli, University of Pennsylvania
- Daniel Huson, University of Tübingen
- Rui Jiang, Tsinghua University
- Tao Jiang, University of California Riverside
- Samuel Kaski, Helsinki University of Technology
- Uri Keich, Cornell University
- Luhua Lai, Peking University
- Tak Wah Lam, University of Hong Kong
- Charles Lawrence, Brown University
- Chris Lee, University of California at Los Angeles
- Doheon Lee, KAIST
- Sang Yup Lee, KAIST
- Cheng Li, Dana-Farber Cancer Institute /Harvard School of Public Health
- Hao Li, University of California at San Francisco
- Jinyan Li, Nanyang Technological University
- Wen-Hsiung Li, University of Chicago
- Wentian Li, Feinstein Institute for Medical Research
- Yixue Li, Chinese Academy of Sciences and Shanghai Jiaotong University
- Jun S. Liu, Harvard University
- Jingchu Luo, Peking University
- Bin Ma, University of Western Ontario
- Hiroshi Mamitsuka, Kyoto University
- Satoru Miyano, University of Tokyo
- Kenta Nakai, University of Tokyo
- Jurg Ott, Beijing Institute of Genomics
- Laxmi Parida, IBM Watson Research Center
- John Quackenbush, Harvard School of Public Health
- Nikolaus Rajewsky, Max Delbrück Center for Molecular Medicine-Berlin

- Ron Shamir, Tel Aviv University
- Steven Skiena, State University of New York at Stony Brook
- Yun S. Song, University of California at Berkeley
- Paul Spellman, Lawrence Berkeley National Laboratory
- Fengzhu Sun, University of Southern California
- Zhirong Sun, Tsinghua University
- Masaru Tomita, Keio University
- Martin Vingron, Max Planck Institute for Molecular Genetics
- Lusheng Wang, City University of Hong Kong
- Limsoon Wong, National University of Singapore
- Wing H. Wong, Stanford University
- Eric Xing, Carnegie Mellon University
- Ying Xu, University of Georgia
- Hong Xue, Hong Kong University of Science and Technology
- Hong Yan, City University of Hong Kong
- Weichuan Yu, Hong Kong University of Science and Technology
- Luoxin Zhang, National University of Singapore
- Hongyu Zhao, Yale University
- Qing Zhou, University of California at Los Angeles

Table of Contents

Category	Subcategory	Pages (papers; posters)
A – DNA sequence processing	A1. Sequence alignment and evolution A2. Sequence assembly	12; 823 141
B – Expression and regulation	B1. Gene expression, microarray data analysis and disease classification B2. Gene regulation	185; 837 325; 844
C – non-coding RNAs	C1. miRNAs and RNAi C2. RNA structure	374; 856 426; 864
D – Proteins and proteomics	D1. Protein structure, location and function D2. Mass spectrometry data processing	464; 867 556; 885
E – Pathways and networks	E1. Pathways E2. Networks and systems biology	588; 889 631; 891
F – Genomic variations and complex diseases	F1. Disease classification and metabolism modeling F2. Association study and genomic variation	673; 908 717; 915

Preface: Introduction to APBC2009

Michael Q. Zhang, Michael S. Waterman, Xuegong Zhang

1

Oral Presentations

Keynote	<i>David J. Lipman:</i> A Molecular Picture of Seasonal Influenza	4
Keynote	<i>Michael S. Waterman:</i> Sequence Analysis Using Eulerian Graphs	5
Invited	<i>Michael B. Eisen:</i> Understanding and Exploiting the Evolution of <i>Drosophila</i> Regulatory Sequences	6
Invited	<i>Bailin Hao:</i> Independent Verification of 16S rRNA Based Prokaryotic Phylogeny By Composition Vector Approach	8
Invited	<i>John Mattick:</i> A New Understanding of the Human Genome	9
Invited	<i>Pavel Pevzner:</i> EULER: From Assembling Short DNA Reads to Shotgun Protein Sequencing by Assembling Mass Spectra	10
Invited	<i>Martin Vingron:</i> A Molecular Picture of seawnd Influenza	11

A – DNA Sequences

A1. Sequence alignment and evolution

A1-01	Genome aliquoting with double cut and join. <i>Robert Warren and David Sankoff.</i>	12
-------	---	----

A1-02	Pathogenic <i>Bacillus anthracis</i> in the progressive gene losses and gains in adaptive evolution. <i>GongXin Yu</i> .	22
A1-03	SE: An algorithm for deriving sequence alignment from a pair of superimposed structures. <i>Chin-Hsien Tai, James J. Vincent, Changhoon Kim and Byungkook Lee</i> .	32
A1-04	ComPhy: prokaryotic composite distance phylogenies inferred from whole-genome gene sets. <i>Guan Ning Lin, Zhipeng Cai, Guohui Lin, Sounak Chakraborty and Dong Xu</i> .	40
A1-05	Maximum independent sets of commuting and noninterfering inversions. <i>Krister Swenson, Yokuki To, Jijun Tang and Bernard Moret</i> .	53
A1-06	Inversion-based genomic signatures. <i>Krister Swenson and Bernard Moret</i> .	63
A1-07	Triplet supertree heuristics for the tree of life. <i>Harris T. Lin, J. Gordon Burleigh and Oliver Eulenstein</i> .	72
A1-08	Sorting by reversals, block interchanges, tandem duplications, and deletions. <i>Martin Bader</i> .	83
A1-09	Improved algorithms for approximate string matching (extended abstract). <i>Dimitris Papamichail and Georgios Papamichail</i> .	94
A1-10	Simultaneous phylogeny reconstruction and multiple sequence alignment. <i>Feng Yue, Jian Shi and Jijun Tang</i> .	105
A1-11	Methods for comparative metagenomics. <i>Daniel Huson, Daniel C Richter, Suparna Mitra, Alexander F Auch and Stephan C Schuster</i> .	115
A1-12	DNA motif alignment by evolving a population of Markov chains. <i>Chengpeng Bi</i> .	126
A1-13	Efficient Estimation of the Accuracy of the Maximum Likelihood Method for Reconstruction of Ancestral States. <i>Bin Ma and Louxin Zhang</i> .	138

A2. Sequence assembly

A2-01	Parallel short sequence assembly of transcriptomes. <i>Benjamin G Jackson, Patrick S. Schnable and Srinivas Aluru</i> .	141
A2-02	Finding optimal threshold for correction error reads in DNA assembling. <i>Francis Y.L. Chin, Henry C.M. Leung, Weilin Li and Siu-Ming Yiu</i> .	153
A2-03	Crystallizing short-read assemblies around seeds. <i>Mohammad Sajjad Hossain, Navid Azimi and Steven Skiena</i> .	162
A2-04	Short read DNA fragment anchoring algorithm. <i>Wendi Wang, Peiheng Zhang, Xinchun Liu</i> .	173

B – Expression and Regulation

B1. Gene expression, microarray data analysis and disease classification

B1-01	REMAS: a new regression model to identify alternative splicing events from exon array data. <i>Hao Zheng, Xingyi Hang, Ji Zhu, Minping Qian, Wubin Qu, Chenggang Zhang, Minghua Deng</i> .	185
B1-02	A voting approach to identify a small number of highly predictive genes using multiple classifiers. <i>Md. Rafiul Hassan, M. Maruf Hossain, James Bailey, Geoff Macintyre, Joshua W. K. Ho and Kotagiri Ramamohanarao</i> .	194

B1-03	Knowledge driven decomposition of tumor expression profiles. <i>M.H. van Vliet, L.F.A. Wessels and M.J.T. Reinders.</i>	206
B1-04	Network-based support vector machine for classification of microarray samples. <i>Yanni Zhu, Xiaotong Shen and Wei Pan.</i>	220
B1-05	Using random forest for reliable classification and cost-sensitive learning for medical diagnosis. <i>Fan Yang, Hua-zhen Wang, Hong Mi, Cheng-de Lin, Wei-wen Cai.</i>	231
B1-06	A statistical framework for integrating two microarray data sets in differential expression analysis. <i>Yinglei Lai, Sarah E. Eckenrode and Jin-Xiong She.</i>	245
B1-07	Comparison of Affymetrix data normalization methods using 6,926 experiments across five array generations. <i>Reija Autio, Sami Kilpinen, Matti Saarela, Olli Kallioniemi, Sampsa Hautaniemi, Jaakko Astola.</i>	257
B1-08	Integrative disease classification based on cross-platform microarray data. <i>Chun-Chi Liu, Jianjun Hu, Mrinal Kalakrishnan, Haiyan Huang, Xianghong Jasmine Zhou.</i>	269
B1-09	Principal component tests: applied to temporal gene expression data. <i>Wensheng Zhang, Hong-bin Fang, Jiuzhou Song.</i>	278
B1-10	Biclustering of gene expression data using reactive greedy randomized adaptive search procedure. <i>Smitha Dharan, Achuthsankar S Nair.</i>	287
B1-11	Mutual information estimation reveals global associations between stimuli and biological processes. <i>Taiji Suzuki, Masashi Sugiyama, Takafumi Kanamori and Jun Sese.</i>	297
B1-12	Human gene expression sensitivity according to large scale meta-analysis. <i>Pei Hao, Siyuan Zheng, Jie Ping, Kang Tu, Christian Gieger, Rui Wang-Sattler, Yang Zhong, Yixue Li.</i>	310
B1-13	Advanced analysis and visualization of gene copy number and expression data. <i>Reija Autio, Matti Saarela, Anna-Kaarina Järvinen, Sampsa Hautaniemi, Jaakko Astola.</i>	317

B2. Gene regulation

B2-01	Identification of histone modifications in biomedical text for supporting epigenomic research. <i>Corinna Kolárik , Roman Klinger and Martin Hofmann-Apitius.</i>	325
B2-02	Searching for bidirectional promoters in Arabidopsis thaliana. <i>Quan Wang, Lin Wan, Dayong Li, Lihuang Zhu, Minping Qian, Minghua Deng.</i>	336
B2-03	ModuleDigger: an itemset mining framework for the detection of cis-regulatory modules. <i>Hong Sun, Tijl De Bie, Valerie Storms, Qiang Fu, Thomas Dhollander, Karen Lemmens, Annemieke Verstuyf, Bart De Moor, Kathleen Marcha.</i>	347
B2-04	Transcriptional interaction-assisted identification of dynamic nucleosome positioning. <i>Zhiming Dai, Xianhua Dai, Qian Xiang, Jihua Feng, Yangyang Deng, Jiang Wang, Caisheng He.</i>	360

C – Non-coding RNAs

C1. miRNAs and RNAi

C1-01	Predicting microRNA targets in time-series microarray experiments via functional data analysis. <i>Brian J Parker, Jiayu Wen.</i>	374
C1-02	A structural interpretation of the effect of GC-content on efficiency of RNA interference. <i>Chi Yu Chan, C. Steven Carmack, Dang D. Long, Anil Maliyekke, Yu</i>	385

Shao, Igor B. Roninson, and Ye Ding.

- C1-03 Computational identification of condition-specific miRNA targets based on gene 392
expression profiles and sequence information. *Je-Gun Joung, Zhangjun Fei.*
- C1-04 HHMMiR: Efficient de novo prediction of microRNAs using hierarchical hidden 401
Markov models. *Sabah Kadri, Veronica Hinman, Panayiotis V. Benos.*
- C1-05 CComputational prediction of novel non-coding RNAs in Arabidopsis thaliana. 413
*Dandan Song, Yang Yang, Bin Yu, Binglian Zheng, Zhidong Deng, Bao-Liang Lu,
Xuemei Chen and Tao Jiang.*

C2. RNA structure

- C2-01 Fine-grained parallel RNAalifold algorithm for RNA secondary structure prediction 426
on FPGA. *Fei Xia, Yong Dou, Xingming Zhou, Xuejun Yang , Jiaqing Xu and Yang
Zhang.*
- C2-02 Prediction of RNA secondary structure with pseudoknots using integer 439
programming. *Unyane Poolsap, Yuki Kato and Tatsuya Akutsu.*
- C2-03 Sequence-structure relations of pseudoknot RNA. *Fenix WD Huang, Linda YM Li and 449
Christian M Reidys.*

D – Proteins and Proteomics

D1. Protein structure, location and function

- D1-01 Reduction of the secondary structure topological space through direct estimation 464
of the contact energy formed by the secondary structures. *Weitao Sun, Jing He.*
- D1-02 Co-evolution positions and rules for antigenic variants of human influenza A/H3N2 474
viruses. *Jhang-Wei Huang, Chwan-Chuen King, Jinn-Moon Yang.*
- D1-03 Predicting disordered regions in proteins using the profiles of amino acid indices. 484
Pengfei Han, Xiuzhen Zhang and Zhi-Ping Feng.
- D1-04 A method to improve protein subcellular localization prediction by integrating 494
various biological data sources. *Thai Quang Tung, Doheon Lee.*
- D1-05 Studying the unfolding process of protein G and protein L under physical property 505
space. *Liling Zhao, Jihua Wang, Xianghua Dou Zanzia Cao.*
- D1-06 Prediction of amyloid fibril-forming segments based on a support vector machine. 513
Jian Tian, Ningfeng Wu, Jun Guo, Yunliu Fan.
- D1-07 Towards comprehensive structural motif mining for better fold annotation. *Yi Jia, 523
Jun Huan, Vincent Buhr, Jintao Zhang and Leonidas N. Carayannopoulos.*
- D1-08 Semi-supervised protein subcellular localization. *Qian Xu, Derek Hao Hu, Hong Xue, 537
Weichuan Yu and Qiang Yang.*
- D1-09 A computational analysis of SARS cysteine proteinase-octapeptide substrate 548
interaction: implication for structure and active site binding mechanism.
*Kongsakda Phakthanakanok, Khanok Ratanakhanokchai, Khin Lay Kyu, Pornthep
Sompornpisut, Aaron Watts, and Surapong Pinitglang.*

D2. Mass spectrometry data processing

- D2-01 Quality assessment of tandem mass spectra using support vector machine (SVM). 556

An-Min Zou, Fang-Xiang Wu and Jia-Rui Ding and Guy G. Poirier.

- D2-02 Efficient discovery of abundant post-translational modifications and spectral pairs 568 using peptide mass and retention time differences. Yan Fu, Wei Jia, Zhuang Lu, Haipeng Wang, Zuofei Yuan, Hao Chi, You Li, Liyun Xiu, Wenping Wang, Chao Liu, Leheng Wang, Ruixiang Sun, Wen Gao, Xiaohong Qian, Si-Min He.
- D2-03 Analysis of mass spectrometry data using sub-spectra. Wouter Meuleman, Judith Y.M.N. Engwegen, Marie-Christine W. Gast, Lodewyk F.A. Wessels and Marcel J.T. Reinders. 579

E – Pathways and Networks

E1. Pathways

- E1-01 Predicting allosteric communication pathways using motion correlation network. 588 Tu-Liang Lin and Guang Song.
- E1-02 Comparative analysis of the JAK/STAT signaling through erythropoietin receptor 599 and thrombopoietin receptor using a systems approach. Hong-Hee Won, Inho Park, Eunjung Lee, Jong-Won Kim and Doheon Lee.
- E1-03 Semi-supervised gene shaving method for predicting low variation biological 607 pathways from genome-wide data. Dongxiao Zhu.
- E1-04 Analysis on relationship between extreme pathways and correlated reaction sets. 619 Yanping Xi, Yi-Ping Phoebe Chen, Ming Cao, Weirong Wang and Fei Wang.

E2. Networks and systems biology

- E2-01 Extract interaction detection methods from the biological literature. Hongning 631 Wang, Minlie Huang and Xiaoyan Zhu.
- E2-02 Finding motif pairs in the interactions between heterogeneous proteins via 644 bootstrapping and boosting. Jisu Kim, Kyungsook Han.
- E2-03 A new graph-based method for pairwise global network alignment. Gunnar W. Klau. 652
- E2-04 GAIA: a gram-based interaction analysis tool - an approach for identifying 662 interacting domains in yeast. Kelvin X. Zhang and B.F. Francis Ouellette.

F – Genetic Variations and Complex Diseases

F1. Disease classification and metabolism modeling

- F1-01 On optimal comparability editing with applications to molecular diagnostics. 673 Sebastian Böecker, Sebastian Briesemeister and Gunnar W. Klau.
- F1-02 Genome-scale analysis to the impact of gene deletion on the metabolism of E. coli : 683 constraint-based simulation approach. Zixiang Xu, Xiao Sun, Shihai Yu.
- F1-03 An unsupervised partition method based on association delineated revised mutual 694 information. Jing Chen, Guangcheng Xi.
- F1-04 A new classification method using array Comparative Genome Hybridization data, 706 based on the concept of Limited Jumping Emerging Patterns. Tomasz Gambin and Krzysztof Walczak.

F2. Association study and genomic variation

- F2-01 Copy-number-variation and copy-number-alteration region detection by 717

cumulative plots. *Wentian Li, Annette Lee and Peter K Gregersen.*

- F2-02 Identifying disease associations via genome-wide association studies. *Wenhui Huang, Pengyuan Wang, Zhen Liu and Liqing Zhang.* 729
- F2-03 Comparison of automated candidate gene prediction systems using genes implicated in type 2 diabetes by genome-wide association studies. *Erdahl T Teber, Jason Y Liu, Sara Ballouz, Diane Fatkin and Merridee A Wouters.* 738
- F2-04 CGTS: a site-clustering graph based tagSNP selection algorithm in genotype data. *Jun Wang, Mao-zu Guo, Chun-yu Wang.* 751
- F2-05 PCA-based population structure inference with generic clustering algorithms. *Chih Lee, Ali Abdool and Chun-Hsi Huang.* 761
- F2-06 A fast algorithm for genome-wide haplotype pattern mining. *Søren Besenbacher, Christian N.S. Pedersen and Thomas Mailund.* 772
- F2-07 Detecting Disease-Associated Genotype Patterns. *Quan Long, Qingrun Zhang and Jurg Ott.* 781
- F2-08 A random forest approach to the detection of epistatic interactions in case-control studies. *Rui Jiang, Wanwan Tang, Xuebing Wu and Wenhui Fu.* 788
- F2-09 Partial correlation analysis indicates causal relationships between GC-content, exon density and recombination rate variation in the human genome. *Jan Freudenberg, Mingyi Wang, Yaning Yang and Wentian Li.* 801
- F2-10 Minimizing recombinations in consensus networks for phylogeographic studies. *Laxmi Parida, Asif Javed, Marta Melé, Francesc Calafell and Jaume Bertranpetit and Genographic Consortium.* 811

Posters

- PA1-01 Evolutionary and genomic analyses of *cee*, a novel but ancient gene involved in development. *Jorge Fernandes, Daniel Macqueen, Hung-Tai Lee and Ian Johnston.* 823
- PA1-03 Heterogeneous Computing for Bioinformatics Applications. *Xiandong Meng and Vipin Chaudhary.* 824
- PA1-04 De Novo Exon Acquisition during Human Evolution. *Yoonsoo Hahn and Byungkook Lee.* 825
- PA1-05 Feasibility Study of Using Massively Parallel Pyrosequencing in Detecting Variation in Dengue Virus Genome. *Kwanrutai Chin-inmanu, Aroonroong Jairungsri, Duangjai Sangsrakru, Sithichoke Tangphatsornruang, Somvong Tragoonrung and Prapat Suriyaphol.* 826
- PA1-07 SE: An algorithm for deriving sequence alignment from a pair of superimposed structures. *Chin-Hsien Tai, Changhoon kim, James J. Vincent and Byungkook (BK) Lee.* 827
- PA1-08 Discovery of Unknown Bacteria: A Metagenomic Analysis. *Divya Suryakumar, Raveesh Chilakapati, Qingzhong Liu and Andrew Sung.* 828
- PA1-09 An annotation system focused on metabolic reconstruction and comparative genomics: characterization of an acidophilic microbial community. *Aida Arcas, Noemi Alvarez, Marina Postigo, Mercedes Moreno, Victor Parro and Manuel J Gómez.* 829
- PA1-10 Multiple Weighted Sequence Alignment for Functional Residue Prediction on Carbohydrate Binding Modules. *Wei-Yao Chou, Wei-I Chou, Tun-Wen Pai, Shu-Chuan* 830

Lin, Ting-Ying Chiang, Chuan Yi Tang and Margaret Dah-Tsyr Chang.

- PA1-11 The evolution rate speed-up event of UCP1 indicates the essential residues for its 831
characteristic high uncoupling activity. *Yaping Liu, Kun Chen, Zhiyuan Lv, Hongxiang
Zheng, Jin Wang and Chenyu Zhang.*
- PA1-12 Comparative Metagenomics with MEGAN 2.0. *Suparna Mitra, Daniel Richter, 832
Alexander Auch, Stephan Schuster and Daniel Huson.*
- PA1-13 ComPhy: Prokaryotic Composite Distance Phylogenies Inferred from Whole-Genome 833
Gene Sets. *Guan Ning Lin, Zhipeng Cai, Guohui Lin, Sounak Chakraborty and Dong Xu.*
- PA1-14 Genome-wide analysis of natural selection in human repetitive DNA elements. 834
Jian-Huan Chen, Qian Xu, Siu-Kin Ng and Hong Xue.
- PA1-15 Genome-wide Identification of ChIP-Seq Enriched Regions Based on a Statistical 835
Model. *Chongzhi Zang, Chen Zeng, Weiqun Peng, Dustin E Schones, Kairong Cui and
Keji Zhao.*
- PA1-16 Sorting by inversions, fast. *Krister Swenson, Vaibhav Rajan, Yu Lin and Bernard 836
Moret.*
- PB1-01 Singular value decomposition for genome-wide spermatogenesis microarray data 837
analysis. *Weixiang Liu, Aifa Tang, Kehong Yuan, Datian Ye, Tianfu Wang, Siping Chen.*
- PB1-02 Gene Selection with Fold Change Detection via Nonnegative Matrix Factorization. 838
Weixiang Liu, Tianfu Wang, Siping Chen, Aifa Tang, Kehong Yuan, Datian Ye.
- PB1-03 A joint mixture model for clustering genes from Gaussian and beta distributed data. 839
Xiaofeng Dai, Timo Erkkilä, Olli Yli-Harja and Harri Lähdesmäki.
- PB1-04 Research on Random Forests-based gene selection of microarray data of colon 840
cancer. *Jiangeng Li, Zhikun Gao, Zhi Yan and Youyong Lu.*
- PB1-05 Gene Set Enrichment Analysis for Replicated Microarray Time Course Data. *Jae-Young 841
Kim, Hyungmin Lee and Miyoung Shin.*
- PB1-08 Data Analysis Tools at the UCM-CNB Bioinfo Unit. *Pedro Carmona-Saez, Ruben 842
Nogales-Cadenas, Miguel Vazquez, Edgardo Mejía, Federico Abascal, Mariana Neves,
Francisco Tirado, Jose-Maria Carazo and Alberto Pascual-Montano.*
- PB2-01 Prevalent use of bidirectional promoters generates pervasive transcription in yeast. 843
Zhenyu Xu, Wu Wei, Julien Gagneur and Lars Steinmetz.
- PB2-02 Genome-wide identification of transcriptional start sites, promoters and 844
transcriptional factor binding sites in Escherichia coli K-12 bacterial genome:
Development of a bioinformatics tool for the graphical display of the high throughput
sequencing data. *Enrique Morett, Alfredo Mendoza, Leticia Olvera, Maricela Olvera,
Katy Juárez, contreras-moreira, Verónica Jiménez, Julio Collado-Vides, Blanca
Taboada, Leticia Vega-Alvarado, Ryosuke Watanabe and Edgar Vallejo.*
- PB2-03 Prediction of translation initiation site for microbial genomes with TriTISA. *Gang-Qing 845
Hu, Xiaobin Zheng, HuaiQiu Zhu and Zhen-Su She.*
- PB2-05 A TaqMan® real-time PCR based method for studying copy number variations. 846
Xiaoqing You, Matt Xia, Kelly Li and Gene Spier.
- PB2-06 Evaluation and expression of "drug-likeness" of small molecules using Semantic Web 847
technologies. *Jooyoung Choi, Melissa Davis and Mark A. Ragan.*

- PB2-07 Chromatin modification similarity and coexpression of neighboring gene pairs in yeast. *Yangyang Deng, Xianhua Dai, Qian Xiang, Caisheng He, Jiang Wang, Jihua Feng and Zhiming Dai.* 848
- PB2-08 A comparison study of cross-platform experimental nucleosome positioning data: a nucleosome statistical distribution domain model in the *S. cerevisiae* genome. *Jihua Feng, Xianhua Dai, Qian Xiang, Zhiming Dai, Jiang Wang, Yangyang Deng and Caisheng He.* 849
- PB2-09 A mutation degree model for identifying transcriptional regulatory elements. *Chang Q Zhang, Jin Wang, Huaiqiu Zhu and Xiang Gao.* 850
- PB2-10 Assessing the apparent stability of conserved sites in plant virus genomes over time. *Linda Zheng, Mark Gibbs and Brendan Rodoni.* 851
- PB2-11 A fast algorithm on detecting spatially locally co-expressed regions between genes from *in situ* hybridization data. *Xiaotu Ma and Wenyuan Li.* 852
- PB2-13 Evolutionary Dynamics of Genomic Regulatory Blocks: HCNEs and HARs, the content and the boundaries. *Xianjun Dong, Altuna Akalin, David Fredman and Boris Lenhard.* 853
- PB2-14 Recovering gene regulations by optimization model. *Yong Wang, Xiang-Sun Zhang and Luonan Chen.* 854
- PC1-01 A Target-Structure-Based Hybridization Model for Prediction of MicroRNA:Target Interactions. *Dang Long, Chi Yu Chan, Molly Hammell, Rosalind Lee, Peter Williams, Victor Ambros, and Ye Ding.* 855
- PC1-02 Characteristic analysis of miRNA precursors in metazoan species. *Xiaobai Zhang, Xiaofeng Song, Huinan Wang and Huanping Zhang.* 856
- PC1-04 The microRNAs' effects on the expression of target genes at mRNA and protein levels. *Yunfei Pei, Xi Wang and Xuegong Zhang.* 857
- PC1-05 Rapid Evolution of Mammalian X-linked Testis microRNAs. *Xuejiang Guo, Bin Su, Zuomin Zhou and Jiahao Sha.* 858
- PC1-06 sRNATarBase: a comprehensive database of bacterial sRNA targets verified by experiments. *Yalin Zhao, Yuan Cao, Xiaomin Ying, Lei Cha, Ligui Wang, NingSheng Shao and Wuju Li.* 859
- PC1-07 Cleavage in MicroRNA Maturation by Drosha. *Jiyuan An, Yongjun Fan, Anthony Beckhouse, Alistair Chalk and Christine Wells.* 860
- PC1-08 miRNA analysis, the design of highly discriminative oligonucleotides with a memetic algorithm. *Niels Tolstrup, Marc Juul Christoffersen, Nana Jacobsen, Ditte Andreasen, Peter K Busk, Lena Hansson, Jesper Salomon, Hanni Willenbrock, Jens Friis-Nielsen and Peter Mouritzen.* 861
- PC2-02 Incorporating Statistical Characters of Amino Acid for the Prediction of RNA-Protein Interface Residue. *Xiujun Gong, Xinmi Liu and Hua Yu.* 862
- PC2-03 Computational evidence of A-to-I RNA editing in the nucleus transcriptome of *Arabidopsis thaliana*. *Pufeng Du and Yanda Li.* 863
- PC2-05 Group-specific 16S rRNA probe design using MCMC methods. *Yibo Wu, Lirong Yan, Hui Liu, Hanchang Sun and Hongwei Xie.* 864
- PD1-04 DomML, a novel ab initio protein domain prediction method based on domain size distribution. *Yann Christinat.* 865

- PD1-05 A plasmid-borne O-antigen in *Brachyspira hyodysenteriae*. Phatthanaphong 866
Wanchanthuek, Karon Ryan, Paula Moolhuijzen, Zayed Albertyn, Babak Shaban,
David J. Hampson, Matthew Bellgard.
- PD1-06 Identifying Conserved Regions for Protein Structure Prediction. Yi Wei, Dongbo Bu 867
and Jishuang Yang.
- PD1-07 Classification RpoS regulon in *Burkholderia pseudomallei* based on proteomic profile 868
information. Yupaporn Osiriphun, Patompon Wongtrakoongate, Sucharat Sanongkiet,
Prapat Suriyaphol and Sumalee Tungpradabkul.
- PD1-09 Evolutionary conservation in multiple faces of protein interaction. Sanguk Kim, Yoon 869
Sup Choi, Jae-Seong Yang and Sung Ho Ryu.
- PD1-11 Uncovering signaling networks based on high-throughput data. Xingming Zhao and 870
Luonan Chen.
- PD1-12 Order and Content of Thymine in Protein Coding Frames of mRNA Sequences. 871
Ekambaram Rajasekaran and Perumal Anandagopu.
- PD1-13 Protein Function-Function Correlation Approach for Protein Functions Prediction. 872
Ka-Lok Ng, Chien-Hung Huang and Jin-Shuei Ciou.
- PD1-14 Improved computational prediction of prokaryotic protein subcellular localization. 873
Nancy Yu, James Wagner, Matthew Laird, Raymond Lo, Martin Ester and Fiona
Brinkman.
- PD1-15 Towards Comprehensive Structural Motif Mining for Better Fold Annotation in the 874
"Twilight Zone" of Sequence Dissimilarity. Yi Jia, Vincent Buhr, Jintao Zhang, Jun
Huan and Leonidas N. Carayannopoulos.
- PD1-17 Protease Substrate Site Predictors Based on Multilevel Substrate Phage Display Data. 875
Ching-Tai Chen, Ei-Wen Yang, Wen-Lian Hsu and An-Suei Yang.
- PD1-19 Detecting protein structure flexibility via rigid structure identification. Yu-Feng 876
Huang, Chia-Jui Yang, Chun-Chin Huang, Chien-Kang Huang, Chun-Yi Huang, Li-Yuan
Chiu, Pei-Chun Ko, Chiun-Yao Chiang.
- PD1-20 Prediction of Protein Nuclear Localization Based on Smoothed Evolutionary 877
Information and Probabilistic Latent Semantic Indexing. Emily Chia-Yu Su, Jia-Ming
Chang, Ting-Yi Sung and WenLian Hsu.
- PD1-21 Computational characterisation of human lamina-associated domains (LADs). Wouter 878
Meuleman, Ludo Pagie, Marcel Reinders, Bas van Steensel and Lodewyk Wessels.
- PD1-22 Evolution of protein-protein interaction networks for plant bZIP transcription 879
factors. Ying He, Grigoris Amoutzias and Yves Van de Peer.
- PD1-23 Glutamate Receptor Ion Channels and their relatives in Bacteria. Mao-Feng Ger, 880
Gloria Rendon and Eric Jakobsson.
- PD1-24 CPSARST & CPDB - an efficient search tool and a comprehensive database of circular 881
permutation in proteins. Wei-Cheng Lo, Chi-Ching Lee, Che-Yu Lee, and Ping-Chiang
Lyu.
- PD2-02 Protein maps: Integration of Physico-Chemical Properties for Functional Annotation 882
of Proteins. Sumeet Dua and Pradeep Chowriappa.
- PD2-03 Physical Encoding Based Bioinformatics Analysis for Exploring Protein Sorting Signals. 883
Fan Zhang and Jianjun Hu.

PD2-04	Protein-protein docking by simulating the process of association based on predicted interface. <i>Bingding Huang, Michael Schroeder and Rebecca Wade.</i>	884
PD2-24	Assessment of contribution of genomic data sets to predicting protein functions. <i>Seok-Ha Ko and Hyunju Lee.</i>	885
PE1-01	Bifurcation Analysis of Oscillation Patterns in an NF- κ B Signal Pathway Model. <i>Hong Yue and Baoyun Lu.</i>	886
PE1-03	Constrained Mining of Minimal Separators with Applications to Gene Perturbation Studies. <i>Ying Jin, Naren Ramakrishnan, Lenwood S. Heath and Richard F. Helm.</i>	887
PE2-02	BIOIR: An approach to public domain resource integration for human protein-protein interaction. <i>Hsueh-Chuan Liu, Carlos Roberto Arias and Von-Wun Soo.</i>	888
PE2-04	Identification of disease genes via network alignment of human interactome and phenome. <i>Xuebing Wu, Qifang Liu and Rui Jiang.</i>	890
PE2-05	The Influenza Virus Resource at the National Center for Biotechnology Information. <i>Yiming Bao, Dmitry Dernovoy, Boris Kiryutin, Leonid Zaslavsky, Tatiana Tatusova, Jim Ostell and David Lipman.</i>	891
PE2-06	Dose mitochondrial protein network in yeast have a structure of the dissemination tree? <i>Yoon Kyeong Lee, Myung Ho Yeo, Tae Ho Kang, Jae Soo Yoo and Hakyong Kim.</i>	892
PE2-07	Dao: a novel programming language for bioinformatics. <i>Limin Fu.</i>	893
PE2-08	A functional canonical correlation analysis approach to reverse engineering of <i>in silico</i> gene regulatory networks. <i>Feng Zeng, Xuebing Wu, Xuegong Zhang and Rui Jiang.</i>	894
PE2-09	Semantic map generation for biological databases. <i>Hoan Nguyen, Nicolas Wicker, David Kieffer and Olivier Poch.</i>	895
PE2-10	Using the Non-negative Matrix Factorization model in Bioinformatics. <i>Alberto Pascual-Montano, Ruben Nogales-Cadenas, Pedro Carmona-Saez and Jose-Maria Carazo.</i>	896
PE2-11	The Prediction of Protein-Protein Interaction Networks in Rice Blast Fungus. <i>He Fei, Yan Zhang, Hao Chen, Ziding Zhang and You-Liang Peng.</i>	897
PE2-12	Modelling uncertainty in transcriptome measurements enhances network component analysis of yeast metabolic cycle. <i>Chunqi Chang, Yeung Sam Hung and Mahesan Niranjan.</i>	898
PE2-14	Integrative Online Tool for Simulation of Steady-state and Dynamic Cellular Networks. <i>Hongseok Yun, Choamun Yun, Sang Yup Lee and Sunwon Park.</i>	899
PE2-15	Reconstruction of a Genome-scale Metabolic Network of <i>Clostridium acetobutylicum</i> ATCC 824 and Flux Balance Analysis Combined with Nonlinear Programming. <i>Hongseok Yun, Joungmin Lee and Sang Yup Lee.</i>	900
PE2-16	From Genome to Life-Style. <i>Zhaocheng Fan, Jiayuan Zhao, Xiaoyan zhu, Yang Zhong, Renyi Liu and Tao Jiang.</i>	901
PE2-17	A systematic approach of modulization in hES cell differentiation. <i>Jin Hyuk Kim, Jin-il Han, Hye Young Kim, Min Jung Kim, Bo Kyung Kim and Yong Sung Lee.</i>	902
PE2-18	Construction of comprehensive gene network for mitochondria. <i>Song Gao, Jie Li and Jin Wang.</i>	903

- PE2-19 Network approach for prioritizing drug candidate on gliomas. *Lisha Li, Ningbo Zhang and Shao Li.* 904
- PF1-01 A Case Study: The Development of a Cancer Clinical Trial Database System Using caCORE SDK. *Haibin Wang, Teresa Stilley, Fadlo Khuri and André Rogatko.* 905
- PF1-02 Gene Prospector: ranking genes for investigation of gene-disease associations and gene-environment interactions. *Wei Yu, Anja Wulf, Tiebin Liu, Muin Khoury and Marta Gwinn.* 906
- PF1-03 Computational analysis of disease-related mutation effects on transcription factor binding. *Kirsti Laurila and Harri Lähdesmäki.* 907
- PF1-04 An Approach to the Recovery of Associations between Protein Domains and Complex Diseases. *Wenhui Wang, Wangshu Zhang, Rui Jiang and Yihui Luan.* 908
- PF1-05 Nonlinear cooperation between p53 and ING1 induces bax expression in mouse thymocyte apoptosis. *Shaojin Duan, Lin Wan, Wenjiang Fu, Hong Pan, Qi Ding, Chang Chen, Peiwei Han, Xiaoyan Zhu, Liying Du, Hongxiao Liu, Yuxia Chen, Ximing Liu, Xiting Yan, Minghua Deng and Minping Qian.* 909
- PF1-07 Knowledge driven decomposition of tumor expression profiles. *Martin van Vliet, Lodewyk F.A. Wessels and Marcel J.T. Reinders.* 910
- PF1-09 Gentrepid: a candidate gene prediction webserver. *Jason Liu, Erdahl Teber, Richard George, Sara Ballouz, Chi N Pang, Boer Xu, Naresh P. Bains, Duncan B. Sparrow, Robyn Otway, Guanglan Guo, Diane Fatkin and Merridee Wouters.* 911
- PF2-01 Extended multiple breakpoint graphs and a fast and exact DCJ median solver for linear genomes. *Andrew Wei Xu.* 912
- PF2-02 Epidemic history of HCV genotype 6 in South East Asia. *Yi Tan and Shui Shan Lee.* 913
- PF2-03 PBEAM: A parallel implementation of BEAM for genome-wide inference of epistatic interactions. *Tao Peng, Pufeng Du and Yanda Li.* 914
- PF2-04 SNPSHuttle: bi-directional scan of SNP arrays to gain accuracy in missing genotype inference. *Christine Sinoquet.* 915
- PF2-06 A CCA-based multi-locus method for association studies on multi-phenotype data. *Wanwan Tang and Yanda Li.* 926
- PF2-08 Natural parameter values for generalized gene adjacency. *Zhenyu Yang and David Sankoff.* 927